

**REDACTED
VERSION OF
EXHIBIT 4**



faegredrinker.com

Jared B. Briant
 Partner
 jared.briant@faegredrinker.com
 +1 303 607 3588 direct

Faegre Drinker Biddle & Reath LLP
 1144 15th Street, Suite 3400
 Denver, Colorado 80202
 +1 303 607 3500 main
 +1 303 607 3600 fax

June 4, 2024

VIA EMAIL

Alter-AI-TT@simplelists.susmangodfrey.com

Re: Microsoft's May 17, 2024 Document Production in Authors Guild v. OpenAI, Inc., No. 1:23-cv-08292-SHS, and Alter v. OpenAI, Inc., No. 1:23-cv-10211-SHS (S.D.N.Y.)

Dear Counsel:

We write in response to your May 24, 2024 letter regarding Microsoft's May 17, 2024 document production in this case.

Collection and Production of Custodial Documents

Although an ESI Order has not been entered by the Court, Microsoft has nonetheless proceeded to collect custodial documents in accordance with its obligations under the Federal Rules. In anticipation of an ESI Order that will, broadly speaking, require disclosure of custodians and search terms, Microsoft proposes that the parties disclose custodians and search terms, with corresponding hit reports, they have used or will use to collect custodial documents. To that end, we can propose as early as June 6, 2024 for a mutual exchange. Let us know if Plaintiffs agree. In the meantime, Microsoft has begun to produce the responsive, nonprivileged custodial documents from its searches and is producing those documents on a rolling basis, as set forth in its responses to Plaintiffs' Requests for Production ("RFPs").

"Training"-Related Discovery

Microsoft stands on its objections and responses to Plaintiffs' RFP Nos. 17, 21–24, 27, 28–29, and 40, as they are not deficient and do not require supplementation. We are not aware of any non-privileged document, already produced or not, that, for example, "relat[es] to the use of commercial works of fiction to train CHATGPT," as requested by RFP No. 17, or that, for example, sheds light on "any manner in which any LLM operated by OPENAI has referenced or accessed any published works of fiction or nonfiction, in whole or in part, in any form, format, or language," as requested by RFP No. 21.

Although there may be references in Microsoft's documents to OpenAI's training activities generally, these references do not mean that Microsoft trained OpenAI's GPT models, that Microsoft knew which works were included in OpenAI's training datasets, or, most importantly here, that any such documents are responsive to any of the RFPs identified in your May 24 letter. Your May 24 letter refers generally to "training related requests" and to Microsoft being

Fiction and Non-Fiction Plaintiffs' Counsel

- 2 -

June 4, 2024

"knowledgeable about OpenAI's training,"¹ but that is not what was requested in the text of Plaintiffs' RFPs. The RFPs that Plaintiffs served sought, for example, documents relating to the use of commercial works of fiction/non-fiction to train CHATGPT. (See, e.g., RFP Nos. 17, 18.) As further examples, certain of the RFPs sought documents regarding Microsoft's, and not OpenAI's, training of ChatGPT. See, e.g., RFP No. 22 ("... YOU accessed commercial works of fiction and/or nonfiction, ... to train CHATGPT"), RFP Nos. 28–29 ("... YOUR use of potentially copyrighted material in training CHATGPT"), and RFP No. 39 ("... YOU purportedly used copyrighted material without permission in training CHATGPT"). Such requests do not encompass all documents regarding OpenAI's training activities generally. And as set forth in Microsoft's responses, since Microsoft did not train ChatGPT, it has no documents responsive to these requests.

Given the above, Microsoft does not believe any supplementation of its responses and objections to RFPs, or interrogatories, is required. Nevertheless, we provide the following information with respect to each of the categories of documents identified on Page 2 of your May 24 letter:

-

Microsoft does not intend to “expedite” production of certain categories of documents; Microsoft has already made multiple productions of documents in this case and is continuing to produce responsive documents on a rolling basis.

Plaintiffs' Question Regarding Metadata

The documents produced on May 17, 2024 were noncustodial documents; no metadata has been omitted from these documents.

¹ Moreover, it is not clear from your May 24 letter what Plaintiffs mean by this, given that it was public knowledge that OpenAI was training large language models as far back as 2018, and that Microsoft was building computing infrastructure to handle OpenAI's training activities.

2

Fiction and Non-Fiction Plaintiffs'
Counsel

- 3 -

June 4, 2024

Plaintiffs' Question Regarding Redactions for Relevance

Microsoft has not applied redactions for relevance to any documents. The black box contained in MSFT_AICPY_000004210 appears in the document as it was collected from Microsoft.

Sincerely,



Jared B. Briant

cc: NYClassActions_Microsoft_OHS@orrick.com